**Case study result of cyclistic bike share company**

**Following the steps of data analysis process.**

**Ask**

* **What is the problem you are trying to solve?** 
  + There are cyclists of two groups, members and casual. The company realised that being a member of the cyclist group would benefit the customers as well as increase profitabilty of the company. Therefore the problem lies on how to convert those casual cyclist to members by making them purchase the annual membership scheam.
* **How can your insights drive business decisions?**
* **By answering the following questions we can reach to the business decisions**
  + Why are riders prefer membership rather the the casual?
  + What is there experince being as a caual cyclist or a member?
  + How can casual riders be converted to members?
  + How can we reach our clients to address service excellence?

**Deliverable**

**A Clear statement of the business task**

Cyclistic company which is engaged in sharing bikes for all wants to grow and increase its profit. To meet its objective it needs to study its customers and their riding experince so as to embrace the casual riders to the memebership program and also how to utilize the degital media to reach its marketing stategy. The director of the marketing and Cyclisitc Executive team are the two major stack holders who are responsible for taking the company forward and upon whom the decsion lies following the result of our analysis.

**Prepare**

* **Where is your data located?** 
  + The data is downloadeble in an csv format.
* **How is the data organized?** 
  + The data is organized in a table format in spreadsheet. it is a long data consits of columns like rid\_id, start and end date\_time, start and end station name,start and end station id, start and end latitude, start and end longtude and member category name.
* **Are there issues with bias or credibility in this data? Does your data ROCCC?**
  + Yes, the data is Reliable, Original from trusted Cite which is Complete but not Current (as it is for learning and practice purpose).
* **How are you addressing licensing, privacy, security, and accessibility?**
  + The data is an open source provided as a case study. There for accesible for public to analyse and show case the result.
  + The data is made available by Motivate Internationla Inc undera license.
  + Data-privacy issues prohibit from using riders’ personally identiﬁable information.
* **How did you verify the data’s integrity?** 
  + No matter from where the data comes it should be checked for its validity, completeness and cleaness. Every measure of data formating, checked for complete data and removing duplicate and null values was taken to ensure integrity.
* **How does it help you answer your question?** 
  + We need to know the ride time taken from one station to the other. Since we have both the intial and destination time, it is posible to calculate the difference and reach to the required result through analysis.
* **Are there any problems with the data?**
  + No problem with the data.

**Deliverable**

**A description of all data sources used**

The data is a historical trip data made available by Motivate International Inc to publice for the purpose of this case study. The dataset is licensed but put in public, valid and appropriate to answer the underlaying business questions. However, data privacy has been put to protect the confidentiality of rider’s personally identifieble information. The data gathers comprehensive dataset of all the locations, bike riders and starting and ending date\_time as well. It was organized in a CSV table format that it was suitable for appling excel formulas.

**Process**

* **What tools are you choosing and why?** 
  + I am using Microsoft Excel. Excel has alot of helpful functins used for data cleaning and organizing and am comfortable in using Excel for processing.
* **Have you ensured your data’s integrity?** 
  + Yes, i have checked for duplicate, for date format and completenes of the data. Therefore, the data is valid, complete and clean.
* **What steps have you taken to ensure that your data is clean?** 
  + Cleaning helps the data to be readeable, accurate and visually appleaing.
  + I took the duplicate function accross the sheet for any duplicate row, but found non.
  + I removed 9663 rows that their result does not add value to the analysis, because either the ride\_length is zero or negative.
  + Removed those rows with incomplete data (destination unknown) 1043 records. May bring inaccurate result in case we need to present with the final location.
* **How can you verify that your data is clean and ready to analyze?**

All the irrlevant data were deleted.  
There was not NULL data, Misspelled words, mistyped numbers, extra spaces, duplicate data.  
The date formats is mm/dd/yyyy.  
The column names are meaningful.  
The data adhers to the business logic.

* **Have you documented your cleaning process so you can review and share those results?**

Yes. The cleaning Process was made using Microsoft Excel application. Some of the tools used for cleaning are.

Remove duplicats.  
using filters. Record with blank rows was deleted.  
ride\_length with value of zero and negative numbers were removed.

**Deliverables**

**Documentation of any cleaning or manipulation of data**

Dirty data is data that's incomplete, incorrect, or irrelevant to the problem you're trying to solve. So making analysis based on this data may result in faulty conclusions. Therefore proper data formating was made on date columns using Format cell, duplicate rows was removed using Remove Duplicate function, Sorted and Filtered using the Sort and Filter functions on the menu, removed extra spaces from the cells using TRIM() and some incomplete and irrelevant records was deleted as well. Hence, data integrity was in ensured for the dataset.

**Analyze**

* **How should you organize your data to perform analysis on it?** 
  + Data organization is very important for a Data analyst. Organizing the dataset into categories and subject areas that you can focus on as you analyze.. Formatting and adjusting the data in a table using the filter and sorting functionality helps the analysis process to identify the errors and easily pick out any inconsistenceis.The data should be sorted and filtered before doing any further analysis
* **Has your data been properly formatted?** 
  + Yes, the date is formatted as mm/dd/yyyy HH:MM:SS, the Id is formatted as numbers, the characters as formatted as text.
* **What surprises did you discover in the data?** 
  + Deducting start\_date from end\_date only subtractes the hours not the dates
* **What trends or relationships did you ﬁnd in the data?** 
  + Both cyclists derives their cycle from Monday to Sunday. The number of memeber cyclist out numbers the casual ones in all days of the week. However, the casual riders prefer Sunday for riding their cycles.
* **How will these insights help answer your business questions?**
  + Both members and casual cyclist have diffeent experince on using the ride.
  + Membership subscribers take 72% percent of the total riders.
  + Only 51 of members use greater than 12 hrs (day) and 60,488 less than 12 hrs.166 of casuals use greater than 12 hrs and 23,330 less than 12 hrs from the total of 84,035 rioders.
  + The number of casual riders who use more 12 hrs are 76.5% . which is 52% greater than those who have membership.

**Deliverable**

**Summary of Analysis**

To help us for our work of analaysing the data it was organized in a Pivote table to show us the relatinship or the trend that the two riders have with repsect to using bikes. It was clearly seen that memeber were using more often than the casuals one. However, there are more numbers of casual riders who use bikes the whole day than the members.Therefore if we have more numbers of casual users who are riding more than one day, it will be better for them to be benefited by joining the memebership program.

**Share**

* **Were you able to answer the question of how annual members and casual riders use Cyclistic bikes diﬀerently?** 
  + Yes, Most of the annual memebers use the cycles for short distances compared to the casual ones. Both preffer riding on Sundays more often than the other days.
* **What story does your data tell?** 
  + The data tells that there are two categories of cycle riders at a company called Cyclistic, which is engaged in givieng a cycle riding service to its customers. Casual and Members. The company needs to convert more of its customers from casual to memebers so as the members would benefit the company also will have its income increased.
* **How do your ﬁndings relate to your original question?** 
  + Our findings reavels that there are different and similar experince regarding bike ride between the casual and member users. It also show us some clue on if the casual riders be memebes so as to be benefited from the memebership. Degital medias, email or mobile app can be used for its reachebility and easily influence our clients.
* **Who is your audience? What is the best way to communicate with them?** 
  + As a memeber of Cyclistic Marketing Analytics Team, we have two audiences with in the company. Lily Moreno, Director of Marketing is responsiple for the development of campaigns and intiatives to promot the bike share program. Cyclistic executive team: an excutive team responsible for approving the recommended marketing program.
* **Can data visualization help you share your ﬁndings?** 
  + Data visualization is the presentation of data or findings in a graphical manner. Some data visualization tools are Excel, Google spreadsheet, power point slide show and Tableau. These tools help us to share our findings as well.
* **Is your presentation accessible to your audience?**
  + The presentation uses contrasts, distingushed foreground and background color is made so as to include to vision impired audiences also. Textual calrification is included in case platform incompatibility occures.
  + The presentation can be accessible to the audience by providing a URL or a direct link to it. There fore we can have access to the presentation through a netwrok or internet.

**Deliverable  
Supporting visualizations and key ﬁndings**

In the visualizatin and presentation phase, the main audiences was mentioned, the result of analysis and the tool used for it was discussed to reflect the findings. The visualisation was put simple and cleare and visually apealing giving accessibility to the top priority.

**Act**

* **What is your final conclusion based on your analysis?**
  + Based on the analysis made, there is a greate number of clients who want to use cyclistic bike service and most of them prefer Sunday for their riding. There are more casual riders who use the bike more than 12 hours. These can open a good opportunity for attracting more casual riders to be members of cyclistic. However, there is a need for dedicated research & steps from analysts to develop a sound analytical ground for better decisin making. Attempts should be made to modify the existing memebership schem to provide better benefits and quality service to its riders.
* **How could your team and business apply your insights?**
  + Our conclusion and recommendations
* **What next steps would you or your stackholders take based on your findings?**
  + Stack holders should design, develop and promot the memebership package to attract more casual riders to be registred. The promotion and memebership should be facilitated and supported using digital technology.
* **Is there addtiional data you could use to extend on your findings?**
  + Yes, it would have been good to collect some data on how the health and status of the bikes are handled after being used by both type of riders.

**Deliverable  
Your top three recommendations based on yur analysis**

1. We can promot more casual users to be members because there are higher numbers of casual riders who use more time than the already registerd members.
2. Prepare and provide an attractive memebership package with discounted price.
3. Implement social medias or develope a mobile app suitable for usage to reach riders and tracking as well.

Note: I could also devised two more recommendations to attract more riders to be registered for the membership.

1. Surprizing the custoers with various gifts including a new bike for those members who frequently rides.
2. Discounted price for those registred more than one person from the same familyhose.